**Official Title:** **PHASE III, RANDOMIZED, DOUBLE BLIND, PLACEBO-CONTROLLED STUDY OF VEMURAFENIB (RO5185426) ADJUVANT THERAPY IN PATIENTS WITH SURGICALLY RESECTED, CUTANEOUS BRAF-MUTANT MELANOMA AT HIGH RISK FOR RECURRENCE**

**NCT Number:** NCT01667419

**Document Dates:** **STATISTICAL ANALYSIS PLAN**

Version 4: 17 March 2017

# STATISTICAL ANALYSIS PLAN

**TITLE:** **A PHASE III, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY OF VEMURAFENIB (RO5185426) ADJUVANT THERAPY IN PATIENTS WITH SURGICALLY RESECTED, CUTANEOUS BRAF-MUTANT MELANOMA AT HIGH RISK FOR RECURRENCE**

**PROTOCOL NUMBER:** GO27826

**STUDY DRUG:** Vemurafenib (RO5185426)

**VERSION NUMBER:** 4

**IND NUMBER:** 73620

**EUDRACT NUMBER:** 2011-004011-24

**SPONSOR:** F. Hoffmann-La Roche Ltd

**PLAN PREPARED BY:** [redacted] Ph.D.

**DATE FINAL:** See electronic date stamp below.

| Name | Reason for Signing | Date and Time (UTC) |
|---|---|---|
| [redacted] | Company Signatory (Clinical) | 17-Mar-2017 16:00:01 |

## RATIONALE FOR AMENDMENT (VERSION 4) TO THE STATISTICAL ANALYSIS PLAN

The Statistical Analysis Plan (SAP) has been amended to change the primary disease-free survival (DFS) analysis for Cohort 2.

The revised plan is to perform the primary DFS analysis for Cohort 2 at approximately 105 events, compared with120 events specified in the SAP Version 3. A total of 105 DFS events provides approximately 80% power to detect a hazard ratio (HR) of 0.58 for the vemurafenib arm versus the placebo arm at an overall 2-sided 0.05 significance level, or approximately 74% power to detect a HR of 0.6 as specified in the previous SAP Version 3. This change now corresponds to an improvement in median DFS from 7.7 months in the placebo arm to 13.3 months in the vemurafenib arm. No change will be made to the analysis plan for Cohort 1. The type 1 error rate would be strictly controlled at the 0.05 level by the continued use of the hierarchical testing procedure as pre-specified in the previous SAP Version 3.

Additional minor changes have been made to improve clarity and consistency.

## TABLE OF CONTENTS

## LIST OF TABLES



## LIST OF APPENDICES

## 1. BACKGROUND

The Study GO27826 was designed as a pivotal study to support the submission of a supplemental New Drug Application for the use of vemurafenib for the adjuvant treatment of patients with surgically resected, cutaneous malignant melanoma at high risk for recurrence. This Statistical Analysis Plan (SAP) provides the planned analyses for Study GO27826. For purposes of registration, the analyses outlined in this SAP will supersede those specified in the protocol.

## 2. STUDY DESIGN

Study GO27826 is a Phase III, international, multicenter, double-blind, randomized, placebo-controlled study in patients with completely resected, *BRAF*$^{V600}$ mutation–positive melanoma, as detected by the **cobas**® BRAF V600 Mutation Test, at high risk for recurrence.

Approximately 475 patients will be enrolled into two separate cohorts:

- Cohort 1 (approximately 300 patients) will include patients with completely resected Stage IIC, IIIA (patients with one or more nodal metastasis $>1$ mm in diameter), or IIIB cutaneous melanoma, as defined by the American Joint Committee on Cancer Classification, Version 7 (Balch et al. 2009).
- Cohort 2 (approximately 175 patients) will include patients with Stage IIIC cutaneous melanoma, as defined by this classification scheme.

Eligible patients will be randomized (1:1 ratio) to receive placebo or vemurafenib over a 52-week period, with randomization stratified by pathologic stage (Stage IIC, Stage IIIA, Stage IIIB) and region (North America, Australia/New Zealand/South Africa/Latin America, rest of the world) in Cohort 1 and by region (North America, Australia/New Zealand/South Africa/Latin America, rest of the world) in Cohort 2.

Within each cohort, patients will receive study treatment according to one of the following treatment arms:

- Arm A: placebo orally, twice daily (BID) for 52 weeks (thirteen 28-day cycles)
- Arm B: vemurafenib 960 mg orally, BID for 52 weeks (thirteen 28-day cycles)

Randomization will occur within 90 days after definitive surgery (i.e., the last surgery required for the treatment or the diagnosis of melanoma), and study treatment will begin within 4 calendar days after randomization.

Crossover to vemurafenib treatment will not be allowed for patients receiving placebo.

The final analysis of the primary endpoint of disease-free survival (DFS) will occur for each cohort after the targeted number of events for each cohort is reached (approximately 120 DFS events for Cohort 1 and 105 DFS events for Cohort 2).

## 2.1 OUTCOME MEASURES

### 2.1.1 Primary Efficacy Outcome Measure

The primary outcome measure for this study is as follows:

- DFS will be defined as the time from randomization until the date of the first local, regional, or distant melanoma recurrence, occurrence of new primary melanoma, or death from any cause. The DFS component of melanoma recurrence will be assessed by the investigator. The DFS component of an occurrence of a new primary melanoma will be based upon the diagnosis made by a Roche-designated central pathology laboratory. See Protocol Section 4.5.1.4 for histopathologic and imaging requirements for documentation of recurrence.

### 2.1.2 Secondary Efficacy Outcome Measures

The secondary outcome measures for this study are as follows:

- Distant metastasis–free survival (DMFS) will be defined as the time from randomization until the date of diagnosis of distant (i.e., non-locoregional) metastases or death from any cause.
- Overall survival (OS) will be defined as the time from randomization to the date of death from any cause.

### 2.1.3 Exploratory Efficacy Outcome Measures

The exploratory outcome measures for this study are as follows:

- Retrospective identification of study patients whose tumors harbor non-E, activating mutations of BRAF kinase at amino acid position 600 (e.g., $BRAF^{V600K}$), with use of DNA sequencing methods as a means to assess clinical outcomes in this patient subgroup
- Levels of candidate tumor biomarkers in plasma and serum (e.g., circulating mutant *BRAF* DNA) at different timepoints during the study compared with baseline as a means to monitor for and predict melanoma recurrence or occurrence of a new primary melanoma
- Candidate tumor biomarkers at the protein, RNA, and DNA levels (including *RAS* mutations) that may characterize the molecular phenotype of tumors at melanoma recurrence or occurrence of a new primary melanoma as well as predict development of resistance to adjuvant vemurafenib treatment

- Molecular characterization of squamous cell carcinoma (SCC; cutaneous [including keratoacanthoma/KA] and non-cutaneous) or other new primary neoplasms that may be observed in patients treated with vemurafenib

### 2.1.4 Pharmacokinetic Efficacy Outcome Measures

The pharmacokinetic (PK) outcome measures for this study are as follows:

- Plasma concentrations of vemurafenib at clinically relevant timepoints, including steady–state trough values as well as those associated with diagnosis of SCC, dose interruption, and/or reduction for toxicity, melanoma recurrence, and occurrence of a new primary melanoma

### 2.1.5 Safety Outcome Measures

The safety outcome measures for this study are as follows:

- Incidence, nature, and severity of adverse events, serious adverse events, and adverse events of special interest. Severity will be graded according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE), v4.0.
- Changes from baseline in ECG findings and targeted clinical laboratory analytes during the course of study treatment

### 2.1.6 Patient-Reported Outcome Measure

The patient-reported outcome (PRO) measure for this study is as follows:

- To assess patient-reported symptoms, functional interference, and health–related quality of life in the vemurafenib and placebo treatment arms with use of European Organisation for Research and Treatment of Cancer 30-item Quality of Life Questionnaire (EORTC QLQ-C30)

## 2.2 DETERMINATION OF SAMPLE SIZE

The overall Type I error ($\alpha$) for this study is 0.05 (two sided). The primary objective of the protocol to assess efficacy as measured by DFS will be evaluated separately for each of the two cohorts. The sample size determination was evaluated separately for each cohort.

## 2.3 SAMPLE SIZE AND ANALYSIS TIMING

The final primary efficacy endpoint (DFS) analyses for the two cohorts will be conducted according to the procedure specified in Section 4.4.1.

### 2.3.1 Cohort 1

The final analysis of the primary endpoint of DFS for Cohort 1 will take place when approximately 120 DFS events have occurred, on the basis of the following assumptions:

- Two-sided, stratified log-rank test at the 0.05 significance level

- 80% power
- Median DFS for the control arm of 24 months and estimated median DFS in the vemurafenib treatment arm of 40 months (which corresponds to a hazard ratio [HR] of 0.60)
- 5% annual loss to follow-up for DFS
- No interim analysis

Assuming an accrual rate of 8 patients per month in Cohort 1 and a 17–month ramp-up period to reach steady-state enrollment, approximately 300 patients will be required to be enrolled in Cohort 1 during 43 months and followed for an additional 7 months in order to observe 120 DFS events.

On the basis of the assumptions above, 120 DFS events are projected to occur in Cohort 1 approximately 50 months after the first patient is randomized in this study. At that time, it is projected that median follow-up time will be 21 months in Cohort 1, and the minimum follow-up time (e.g., for the last patient randomized) is projected to be 7 months. Also on the basis of the assumptions of 120 DFS events required and a target HR of 0.60, it is projected that an observed HR of 0.70 or better in the DFS analysis will result in a statistically significant difference between treatment arms (i.e., HR of 0.70 is the minimally detectable difference for that analysis).

A summary of the assumptions and characteristics of the DFS analysis for Cohort 1 is shown in Table 1.

For Cohort 1, on the basis of the assumptions of a target HR of 0.60, 120 DFS events would provide 80% power to detect a 16% absolute increase in the 2-year DFS rate (50% vs. 66%, corresponding to a 40% risk reduction; i.e., HR of 0.60).

### 2.3.2 Cohort 2

The final analysis of the primary endpoint of DFS for Cohort 2 will take place when approximately 105 DFS events have occurred, on the basis of the following assumptions:

- Two-sided, stratified log-rank test at the 0.05 significance level
- 80% power
- Median DFS for the control arm of 7.7 months and estimated median DFS in the vemurafenib treatment arm of 13.3 months (which corresponds to an HR of 0.58)
- 5% annual loss to follow-up for DFS
- No interim analysis

Assuming an accrual rate of 5 patients per month in Cohort 2 and a 27−month ramp-up period to reach steady-state enrollment, approximately 175 patients will be required to be enrolled in Cohort 2 over 40 months in order to observe 105 DFS events.

A summary of the assumptions and characteristics of the DFS analysis for Cohort 2 is shown in Table 1.

For Cohort 2, on the basis of the assumptions of a target HR of 0.58, 105 DFS events would provide 80% power to detect a 17% absolute increase in the 2-year DFS rate (12% vs. 29%, corresponding to a 42% risk reduction; i.e., HR of 0.58).

**Table 1 Assumptions and Characteristics for Disease-Free Survival Analyses by Cohort**

| | Cohort 1 | Cohort 2 |
|---|---|---|
| Patients enrolled | 300 | 175 |
| Hazard ratio targeted | 0.60 | 0.58 |
| Target median (control) | 24 months | 7.7 months |
| Target median (vemurafenib) | 40 months | 13.3 months |
| Final DFS analysis | | |
| Number of DFS events | 120 | 105 |
| MDD hazard ratio [a] | 0.70 | 0.68 |
| $\alpha$ level (two sided) | 0.05 | 0.05 |
| Power | 80% | 80% |

DFS = disease-free survival; FPI = first patient in; MDD = minimum detectable difference.

Note: A 5% annual dropout rate is anticipated for DFS analyses.

[a] Minimally detectable difference; the largest observed hazard ratio that is projected to be statistically significant.

# 3. STUDY CONDUCT

## 3.1 RANDOMIZATION ISSUES

After written informed consent has been obtained and eligibility has been established, each patient will be assigned an identification number and randomized to one of the two treatment arms with the use of an interactive voice or Web response system (IxRS). As noted in Section 2, randomization will be stratified by pathologic stage (Stage IIC, Stage IIIA, Stage IIIB) and region (North America, Australia/New Zealand/South Africa/Latin America, rest of the world) in Cohort 1 and by region (North America, Australia/New Zealand/South Africa/Latin America, rest of the world) in Cohort 2. A stratified, permuted, block randomization scheme will be used to obtain approximately a 1:1 allocation between the two treatment groups.

## 3.2 INDEPENDENT REVIEW FACILITY

An independent review facility will not be used for this study.

## 3.3 DATA MONITORING

An independent Data Safety Monitoring Board (DSMB) will be employed to evaluate safety data from this study as specified in the DSMB Charter.

# 4. STATISTICAL METHODS

Descriptive summaries of continuous data for each cohort will include the mean, standard deviation, median, minimum, and maximum, and number of patients. Descriptive summaries of discrete data for each cohort will include the number of patients and incidence as a frequency and percentage.

The baseline value of any variable will be defined as the last available value prior to the first administration of study treatment.

## 4.1 ANALYSIS POPULATIONS

### 4.1.1 Intent-to-Treat Population

The intent-to-treat (ITT) population is defined as all randomized patients, whether or not study treatment was received. The ITT population will be analyzed according to the treatment assigned at randomization.

### 4.1.2 Safety Population

The safety population will include all patients who receive at least one dose of study treatment. Patients who receive at least one dose of vemurafenib will be included in the vemurafenib safety population. The safety population will be analyzed according to the patients' safety group.

### 4.1.3 Pharmacokinetic-Evaluable Population

The PK-evaluable population will include all patients who have received at least one dose of vemurafenib and have provided valid PK assessments. The PK-evaluable population will be analyzed according to the treatment received. The PK population at specific timepoints will vary, depending on the availability of results at confirmed dosing and PK assessment times.

## 4.2 ANALYSIS OF STUDY CONDUCT

Enrollment, eligibility violations, and patient disposition will be summarized for randomized patients by treatment arm. The summary of patient disposition will include whether treatment was completed or discontinued prematurely and the reason for premature treatment discontinuation. Study treatment administration will be summarized by treatment arm for all treated patients.

## 4.3 ANALYSIS OF TREATMENT GROUP COMPARABILITY

Demographic variables, stratification factors, and other baseline characteristics will be summarized for the ITT population by cohorts.

## 4.4 EFFICACY ANALYSIS

Unless otherwise specified, all efficacy analyses will be performed on the ITT population. The primary and all secondary objectives of this study will be evaluated separately for each cohort.

### 4.4.1 Primary Efficacy Endpoint

The primary endpoint, DFS, is defined as the time from randomization until the date of the first local, regional, or distant melanoma recurrence, occurrence of new primary melanoma, or death from any cause. The DFS component of melanoma recurrence will be assessed by the investigator. The DFS component of an occurrence of a new primary melanoma will be based upon the diagnosis made by a Roche-designated central pathology laboratory. For patients without a DFS event at the time of data cutoff, data will be censored at the date of the last disease assessment. For patients who are treated with any non–protocol anti-cancer therapy (defined as systemic therapy or radiation therapy) without a DFS event, the data will be censored at the date of the last disease assessment. For patients without a DFS event for whom no post-randomization disease assessment is available, the data will be censored on the randomization date.

For patients whose recurrence has been proven histologically, the date of melanoma recurrence will be defined as the earliest date of the scan or clinical examination that prompted the biopsy. For patients whose suspicious lesions were deemed not amenable to biopsy (see Protocol Section 4.5.1.4) or for patients who refuse a biopsy, the date of melanoma recurrence will be defined as the earliest date of the scan or clinical examination that would have prompted a biopsy. For patients with an occurrence of a new primary melanoma, the date of the new primary melanoma will be defined as the earliest date of the clinical examination or scan that prompted the biopsy.

The duration of DFS will be calculated as the earliest DFS event date or censoring date minus the randomization date plus 1 day, converted to months.

The final analysis of the primary endpoint of DFS will take place when approximately 120 DFS events have occurred for Cohort 1 and approximately 105 DFS events have occurred for Cohort 2 (see Section 2.3). The final DFS analyses for both cohorts will be conducted at the same time by using the dataset from the same data cutoff date for both cohorts. The primary efficacy analyses will be comparisons of the two treatment groups, using a two-sided,

stratified log-rank test for Cohort 1 and Cohort 2 separately. The statistical significance of the two DFS tests is defined by the rule outlined below.

The hierarchical testing procedure, whereby Cohort 2 is tested first, will be used to adjust for comparison of two treatment groups (vemurafenib vs. control) for the two cohorts, in order to maintain an overall Type I error rate of 0.05 (two sided) for the final analysis of the primary endpoint of DFS. If the p-value for Cohort 2 is $\leq 0.05$, i.e., the vemurafenib treatment group is statistically significantly different from the control group in Cohort 2, then, Cohort 1 will be tested. If the p-value for Cohort 1 is $\leq 0.05$, then the vemurafenib treatment group is also statistically significantly different from the control group in Cohort 1. That is:

- Step 1: If p2 $\leq 0.05$, then vemurafenib treatment group is statistically significantly different from the control group in Cohort 2. Go to step 2.
- Step 2: If p1 $\leq 0.05$, then vemurafenib treatment group is also statistically significantly different from the control group in Cohort 1.

p1 and p2 are the p-values of final DFS analyses in Cohort 1 and Cohort 2, respectively.

Median DFS time will be estimated using the Kaplan−Meier method, and the two-sided 95% CI will be calculated using the method of Brookmeyer and Crowley (1982) for each cohort. In addition, Kaplan-Meier methodology will be used to estimate landmark (e.g., 1-year, 2-year, and 3-year) DFS rates and the associated two-sided 95% CIs for each treatment arm, and the Kaplan-Meier curves will be provided. The HR for DFS (recurrence, new primary melanoma, or death) and the associated two-sided 95% CI will be computed using a stratified Cox proportional hazards model.

The stratification factors in the stratified analyses are the stratification factors used in randomization of patients in each cohort. For Cohort 1, stratified analyses will incorporate two stratification factors: pathologic stage (Stage IIC, Stage IIIA, and Stage IIIB) and region (North America, Australia/New Zealand/South Africa/Latin America, rest of the world). For Cohort 2, stratified analyses will incorporate one stratification factor, region (North America, Australia/New Zealand/South Africa/Latin America, rest of the world).

As a sensitivity analysis, an unstratified log-rank test will be performed and the unstratified HR will be provided for each cohort.

In addition, the following sensitivity analyses will be performed for DFS by study cohort:

- Missing assessments for patients later diagnosed with a DFS event: For patients with a DFS event (other than death) who missed two or more scheduled assessments immediately prior to the event, the date of the event will be replaced by the date of the last disease assessment, plus 1 day. For patients whose DFS event was death and who died after missing two or more scheduled assessments, data will be censored at the date of the last disease assessment.
- Off-schedule assessments with a DFS event: For patients with a DFS event at an off-schedule visit, the date of event will be replaced by the date of the next scheduled disease-assessment visit.

As an exploratory analysis, DFS analyses based on the pooled data from both cohorts will also be performed for descriptive purposes to characterize the benefit of vemurafenib in the total study population. This exploratory analysis will be stratified by region.

### 4.4.2 Secondary Efficacy Endpoints

#### 4.4.2.1 Distant Metastasis–Free Survival

DMFS is defined as the time from randomization until the date of diagnosis of distant (i.e., non-locoregional) metastasis or death from any cause. For patients without a DMFS event at the time of data cutoff, the data will be censored at the date of the last disease assessment. If no post-randomization disease assessment is available, the observation will be censored on the randomization date.

DMFS will be analyzed at the time of the final DFS analysis in each cohort. The analysis methods to be employed for DMFS are the same as those described for the primary endpoint of DFS.

#### 4.4.2.2 Overall Survival

OS is defined as the time from randomization until the date of death from any cause. For patients still alive at the time of analysis, the data will be censored at the date the patient was last known to be alive. If no post-randomization disease assessment is available, the data will be censored on the randomization date. The duration of OS will be calculated as the date of death or censoring date minus the randomization date plus 1 day, converted to months.

The study is not powered for OS, so adequate power statistical testing for this endpoint is not possible. However, some standard OS estimates will be provided by using the same analysis methods as those described for the primary endpoint of DFS.

The final OS analysis for Cohorts 1 and 2 will be performed after the occurrence of approximately 107 and 118 deaths, respectively (projected to occur at approximately Month 72 in each cohort) or at Month 72, whichever occurs first. One interim analysis of OS is planned in each of the two cohorts at the time of the final DFS analysis for both cohorts (see Section 4.9).

OS will be compared between the two treatment arms using a two–sided stratified log-rank test at an overall two-sided 0.05 significance level for Cohort 1 and Cohort 2 separately. The HR for death will be estimated using a stratified Cox model. Two-sided 95% CIs for the HR will be provided.

Stratified analyses will incorporate the same stratification factors for the analysis of DFS.

Kaplan-Meier methodology will be used to estimate median OS and landmark (e.g., 1-year, 2-year, and 3-year) OS rates and the associated two-sided 95% CIs for each treatment arm, and the Kaplan-Meier curves will be provided.

As a sensitivity analysis, an unstratified log-rank test will be performed and the unstratified HR will be provided.

As an exploratory analysis, OS analyses on the basis of the pooled data from both cohorts will also be performed for descriptive purposes to characterize the benefit of vemurafenib in the total study population.

#### 4.4.2.3 Patient-Reported Outcomes

Quality of life (QoL), as measured by EORTC QLQ-C30, will be evaluated for patients with a baseline assessment and at least one post-baseline QLQ-C30 assessment that generate a score. Total QLQ-C30, each domain score (i.e., physical functioning, role functioning, emotional functioning, cognitive functioning, and social functioning), symptom scales, and their changes from baseline will be examined for each timepoint with use of descriptive statistics, including mean, median, standard deviation, and range.

Compliance rates for patients who complete PRO assessments will be assessed over time and by treatment arm.

See Appendix 1 for a description of the calculation of each domain score and the total QLQ-C30.

#### 4.4.2.4 Type I Error Control for the Secondary Efficacy Endpoints

The following procedure will be performed separately for each cohort, provided that the cohort is statistically significant in the primary DFS analysis (according to the rules described in Section 4.4.1).

To control the Type I error rate at the 0.05 level (two sided) for the secondary efficacy endpoint tests, a hierarchical testing approach will be employed to evaluate the statistical significance of the secondary endpoints of DMFS and OS. If the primary DFS analysis meets statistical significance in either cohort, DMFS and then OS will be evaluated for statistical significance at the 0.05 level (two sided) for that cohort.

The secondary endpoint of DMFS will be compared between treatment arms using a two-sided stratified log-rank test at an overall two-sided 0.05 significance level for Cohorts 1 and 2 separately. For each cohort, if the DFS comparison is positive in favor of vemurafenib, the comparison of DMFS between the placebo and vemurafenib arms will be tested at an overall two-sided 0.05 significance level for each cohort. Gated on the successful testing of the comparison of the secondary endpoint, DMFS, in each cohort, the secondary endpoint of OS will be compared between treatment arms using a two-sided stratified log-rank test at an overall two-sided 0.05 significance level for Cohorts 1 and 2 separately. The Lan-DeMets implementation (Lan and DeMets 1983) of the O'Brien-Fleming use function will be used to control the overall Type I error for the OS comparison in each cohort at a two-sided 0.05 significance level. See Section 4.9.1 for details.

### 4.4.3 Subgroup Analyses

The consistency of the treatment effect of vemurafenib on DFS, OS, and DMFS across subgroups defined by demographic and baseline characteristics and stratification factors will be examined within each cohort. Because some subgroups may have small sample sizes, these analyses will be considered exploratory.

The subgroups to be considered include, but are not limited to, the following:

- Disease stage (Stage IIC, Stage IIIA, and Stage IIIB) for Cohort 1
- Age ($\leq$50 years, $>$50 years) at randomization
- Age ($\leq$65 years, $>$65 years) at randomization
- Tumor ulceration (or mitosis $\geq$1/mm$^2$ for T1 lesions) present or not present at randomization
- Race (non-White, White)
- Sex (female, male)
- Region (North America, Australia/New Zealand/South Africa/Latin America, rest of the world)
- Eastern Cooperative Oncology Group Performance Status at randomization (0 or 1)
- Newly diagnosed versus first metachronous recurrence
- Lymph node types: macroscopic, microscopic, and N3 at randomization

- *BRAF* mutation status such as V600E versus non-V600E at randomization

For DFS, OS, and DMFS, the Kaplan-Meier estimated median time will be summarized by treatment arm for each of the subgroups defined above along with a HR (treatment:control) estimated by unstratified Cox regression that is displayed as a Forest plot (Lewis and Clarke 2001).

## 4.5 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

### 4.5.1 Pharmacokinetic Analyses

Descriptive statistics will be used to perform the analysis of plasma concentrations of vemurafenib at clinically relevant timepoints. These timepoints will include all available Cycle 1 data and pre-dose values from all available cycles. In addition, summary statistics may be provided for PK data from patients following the diagnosis of melanoma recurrence or occurrence of a new primary melanoma during study treatment, at the time of diagnosis of SCC (cutaneous [including KA] and non-cutaneous), and at the occurrence of dose-limiting toxicity and concomitant decision to reduce the dose or interrupt or discontinue treatment. The descriptive statistics will include arithmetic means, standard deviations, geometric means, coefficients of variation, medians, and ranges.

Plasma concentrations that are below the limit of quantification for the assay will be replaced with zero for descriptive statistics. Missing PK data points will not be replaced. Pre-dose values that are determined to be measured post-dose will not be included in the analysis.

## 4.6 EXPLORATORY ANALYSES

Descriptive statistics will be provided for the exploratory biomarker analyses.

## 4.7 SAFETY ANALYSES

All safety analyses will be performed on the safety population, unless specified otherwise. All safety data will be summarized using descriptive statistics.

### 4.7.1 Exposure of Study Medication

Exposure to study treatment will be summarized, with the use of descriptive statistics, on the following:

- Length of time on treatment
- Number of days dosed
- Cumulative dose
- Dose intensity (%) (defined as total amount of study treatment received relative to total amount of study treatment expected between the first and last dose)

Dose modification (dose reduction or interruption) will be summarized as follows:

- n (%) of patients with any dose modification (reduction or interruption)

Dose reduction will be summarized as follows:

- n (%) of patients with at least one dose reduction
- Number of dose reductions per patient (mean, median, and range)
- Final dose level after reduction
- Reasons for change in dose

Dose interruptions will be summarized as follows:

- n (%) of patients with at least one dose interruption
- Number of interruptions per patient (mean, median, and range)

### 4.7.2 Adverse Events

Safety analyses will include all patients who receive any amount of study treatment (vemurafenib or placebo). Patients who receive any amount of vemurafenib will be summarized in the vemurafenib treatment arm. Safety will be assessed through summaries of all adverse events, including serious adverse events, adverse events of special interest, and adverse events leading to discontinuation of vemurafenib or placebo. All verbatim descriptions of treatment–emergent adverse events will be mapped to MedDRA thesaurus terms and graded according to the NCI CTCAE v4.0.

The following safety parameters will be summarized by treatment arm for patients in Cohort 1 and Cohort 2 separately as well as for all study patients pooled:

- All adverse events
- All adverse events leading to discontinuation of study treatment
- All serious adverse events
- Adverse events of interest (e.g., cutaneous SCC [including KA], QTc prolongation)
- NCI CTCAE Grade 3 or greater adverse events
- Adverse events resulting in death
- All deaths

### 4.7.3 Deaths

The following summaries of patient deaths will be provided:

- All deaths by primary cause of death
- Incidence and cause of deaths within 30 days of the start of treatment

### 4.7.4 Laboratory Data

Laboratory toxicities will be defined on the basis of local laboratory normal ranges and the NCI CTCAE v4.0. Laboratory test results and normal ranges will be converted from local lab to Standard International units for analysis purposes. For each laboratory parameter, the toxicity grade at baseline and the worst toxicity grade during the treatment period will be summarized by treatment arm.

### 4.7.5 Vital Signs

No analyses of vital signs or physical-examination findings are planned.

## 4.8 MISSING DATA

For DFS, data from patients who are lost to follow-up without DFS event will be censored at the date of the last disease assessment. If no post-randomization disease assessment is available, the data will be censored on the randomization date.

For DMFS, data from patients who are lost to follow-up without documented distant melanoma recurrence will censored at the date of the last disease assessment. The data of the patients without post-randomization disease assessments will be censored on date of randomization.

For OS, data from patients who are lost to follow-up will be analyzed as censored observations on the date the patient was last known to be alive. The data of patients without post-randomization disease assessments will be censored on date of randomization.

## 4.9 INTERIM ANALYSES

An independent DSMB will monitor the safety of patients and will meet periodically to review summaries of selected safety data prepared by the independent Data Coordinating Center. The detailed interim safety analysis plan and the role and responsibilities of the DSMB members are described in the interim analysis plan and separate charter for the DSMB, respectively.

Interim analyses of OS will be performed by the Sponsor (see Section 4.4.2.2).

### 4.9.1 Efficacy

No interim analyses of the primary endpoint, DFS, will be performed.

Two OS analyses (one interim analysis and one final analysis) are planned for each cohort. The OS interim analysis will be performed at the time of the final DFS analysis for both cohorts. The final OS analysis for Cohorts 1 and 2 will be performed after the occurrence of approximately 107 and 118 deaths, respectively (projected to occur at approximately Month 72 in each cohort) or at Month 72, whichever occurs first.

The Lan-DeMets implementation (Lan and Demets 1983) of the O'Brien-Fleming use function will be used to control the overall Type I error for the OS comparison in each cohort at a two-sided 0.05 significance level. Table 2 summarizes the assumptions and characteristics of the analyses for OS.

**Table 2 Assumptions and Characteristics for Overall Survival Analyses by Cohort**

| | Cohort 1<br>n = 300 | Cohort 2<br>n = 175 |
|---|---|---|
| HR targeted | 0.70 | 0.70 |
| Targeted median (control) | 61 months | 24.2 months |
| Targeted median (vemurafenib) | 87.1 months | 34.6 months |
| Interim OS<br>(to be performed at time of final DFS analysis) | | |
| Projected number of events (% of final events) | 64 (60%) | 59 (50%) |
| Final OS | | |
| Number of events (% of final events) | 107 (100%) | 118 (100%) |
| Estimated cutoff date [a] | 72 months after FPI | 72 months after FPI |
| $\alpha$ level (two sided) | 0.05 | 0.05 |
| Power | 0.45 | 0.49 |

DFS = disease-free survival; FPI = first patient in; HR = hazard ratio; OS = overall survival.

Note: 1% annual dropout rate is anticipated for OS analyses.

[a] Estimated data cutoff time from study enrollment date.

## 5. REFERENCES


Balch CM, Gershenwald JE, Soong SJ, et al. Final version of 2009 AJCC melanoma staging and classification. J Clin Oncol 2009;27:6199–206.

Brookmeyer R, Crowley JJ. A confidence interval for median survival time. Biometrics 1982;38:29−41.

Lan KKG and DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika 1983;70;659–63.

Lewis S, Clarke M. Forest plots: trying to see the wood and the trees. BMJ 2001;322:1479–80.

## Appendix 1
## European Organisation for Research and Treatment of Cancer 30-Item Quality of Life Questionnaire (Version 3)

### EORTC QLQ-C30 (version 3)

We are interested in some things about you and your health. Please answer all of the questions yourself by circling the number that best applies to you. There are no "right" or "wrong" answers. The information that you provide will remain strictly confidential.

| | | Not at All | A Little | Quite a Bit | Very Much |
|---|---|---|---|---|---|
| 1. | Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase? | 1 | 2 | 3 | 4 |
| 2. | Do you have any trouble taking a long walk? | 1 | 2 | 3 | 4 |
| 3. | Do you have any trouble taking a short walk outside of the house? | 1 | 2 | 3 | 4 |
| 4. | Do you need to stay in bed or a chair during the day? | 1 | 2 | 3 | 4 |
| 5. | Do you need help with eating, dressing, washing yourself or using the toilet? | 1 | 2 | 3 | 4 |

### During the past week:

| | | Not at All | A Little | Quite a Bit | Very Much |
|---|---|---|---|---|---|
| 6. | Were you limited in doing either your work or other daily activities? | 1 | 2 | 3 | 4 |
| 7. | Were you limited in pursuing your hobbies or other leisure time activities? | 1 | 2 | 3 | 4 |
| 8. | Were you short of breath? | 1 | 2 | 3 | 4 |
| 9. | Have you had pain? | 1 | 2 | 3 | 4 |
| 10. | Did you need to rest? | 1 | 2 | 3 | 4 |
| 11. | Have you had trouble sleeping? | 1 | 2 | 3 | 4 |
| 12. | Have you felt weak? | 1 | 2 | 3 | 4 |
| 13. | Have you lacked appetite? | 1 | 2 | 3 | 4 |
| 14. | Have you felt nauseated? | 1 | 2 | 3 | 4 |
| 15. | Have you vomited? | 1 | 2 | 3 | 4 |
| 16. | Have you been constipated? | 1 | 2 | 3 | 4 |

Please go on to the next page

## Appendix 1
## European Organisation for Research and Treatment of Cancer 30-Item Quality of Life Questionnaire (Version 3) (cont.)

| **During the past week:** | **Not at All** | **A Little** | **Quite a Bit** | **Very Much** |
|---|---|---|---|---|
| 17. Have you had diarrhea? | 1 | 2 | 3 | 4 |
| 18. Were you tired? | 1 | 2 | 3 | 4 |
| 19. Did pain interfere with your daily activities? | 1 | 2 | 3 | 4 |
| 20. Have you had difficulty in concentrating on things, like reading a newspaper or watching television? | 1 | 2 | 3 | 4 |
| 21. Did you feel tense? | 1 | 2 | 3 | 4 |
| 22. Did you worry? | 1 | 2 | 3 | 4 |
| 23. Did you feel irritable? | 1 | 2 | 3 | 4 |
| 24. Did you feel depressed? | 1 | 2 | 3 | 4 |
| 25. Have you had difficulty remembering things? | 1 | 2 | 3 | 4 |
| 26. Has your physical condition or medical treatment interfered with your family life? | 1 | 2 | 3 | 4 |
| 27. Has your physical condition or medical treatment interfered with your social activities? | 1 | 2 | 3 | 4 |
| 28. Has your physical condition or medical treatment caused you financial difficulties? | 1 | 2 | 3 | 4 |

For the following questions please circle the number between 1 and 7 that best applies to you

29. How would you rate your overall health during the past week?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Very poor | | | | | | Excellent |

30. How would you rate your overall quality of life during the past week?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Very poor | | | | | | Excellent |